

CRISP-DM

Cross-Industry Standard Process for Data Mining

Proporciona una descripción normalizada del ciclo de vida de un proyecto estándar de análisis de datos y consiste en 6 fases

Metodología para desarrollar un proyecto:

Permite el desarrollo y manejo de un proyecto y puedan ser replicados. Cubre la fases de un proyecto, sus tareas y las relaciones entre ellas.

Ayuda a la planeación y gerencia del proyecto.

Contempla el proceso de análisis de datos.

No se interpone o sustituye otras metodologías de desarrollo de software como agile

presión del negocio

Entendimiento de los objetivos y requerimientos del proyecto

Definición del problema

Preparación de Datos

Selección de datos

Limpieza de datos

Evaluación

Determinar si los resultados coinciden con los objetivos del negocio

Identificar los temas de negocio que deberían haberse abordado

Despliegue

CRISP-DM Fases

Compresión de los datos

Obtención del conjunto inicial de datos Exploración del conjunto de datos
Identificar s característica de calidad de los datos Identificar los resultados iniciales
obvios

Modelamiento

Implementación de herramientas de minería de datos

Desplegar nuestro modelo o herramientas dentro de un ambiente como Kubernetes